ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-6327-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICASSP49357.2023.10097091

TOWARDS INTERPRETABLE SEIZURE DETECTION USING WEARABLES

Irfan Al-Hussaini ⋆, *Cassie S. Mitchell* †

Georgia Institute of Technology
⋆alhussaini.irfan@gatech.edu, †cassie.mitchell@bme.gatech.edu

ABSTRACT

Seizure detection using machine learning is a critical problem for the timely intervention and management of epilepsy. We propose `SeizFt`, a robust seizure detection framework using EEG from a wearable device. It uses features paired with an ensemble of trees, thus enabling further interpretation of the model's results. The efficacy of the underlying augmentation and class-balancing strategy is also demonstrated. This study was performed for the Seizure Detection Challenge 2023, an ICASSP Grand Challenge.

***Index Terms*—** seizure, eeg, augmentation, xai, interpretability, imbalanced classes, electroencephalogram

1. INTRODUCTION

The analysis of EEG recordings by domain experts to detect seizures is arduous and expensive but essential for the diagnosis of epilepsy. Therefore, the development of an automated wearable EEG-based epileptic seizure detection system would be invaluable for people with epilepsy [1]. The detection of seizures is an imbalanced binary classification problem, where seizure and non-seizure are the two labels. Recently, deep learning approaches have shown great promise in seizure detection.

We propose `SeizFt`, a seizure detection framework with data augmentation and class balancing strategies buttressing robust features for an interpretable model using an ensemble of trees. Despite the challenges in designing and implementing a feature-based interpretable model, it enhances trust and accountability and enables diagnosis of model performance [2]. This work was performed for the Seizure Detection Challenge 2023[1], an ICASSP Signal Processing Grand Challenge.

The Seizure Detection Challenge aimed to develop machine learning frameworks for accurately detecting seizures in patients with epilepsy using EEG data obtained from a discreet wearable device with behind-the-ear electrodes. The challenge focused on two tasks.

Task 1 (T1) of the challenge involved developing a machine learning model for detecting seizures in wearable SensorDot (SD) data using the entire SeizeIT1 [3] dataset with wearable EEG data from the SD device, single-channel ECG data, and the full scalp EEG data. `SeizFt` obtains 56.88 Total Points on the test set, SeizeIT2 [1].

Task 2 (T2) focused on applying data manipulation techniques to obtain the best performance of a provided Deep Learning model for wearable seizure detection. The model is an adapted version of ChronoNet [4] and had to be trained using the same dataset as T1 [3]. The proposed approach for T2 obtains 18.30 Total Points on the test set, SeizeIT2 [1].

2. METHOD

We propose `SeizFt`, an interpretable feature-based approach to seizure detection for T1. It comprises the following sequence of steps during training:

1. Fourier Transform (FT) Surrogates [5, 6] is used to augment the EEG signals only during training and balance the number of seizure and non-seizure epochs.
2. Robust features are extracted from the EEG, inspired by recent work in interpretable sleep staging [7, 8]. Features include the following: Standard Deviation (STD), Inter-Quartile Range (IQR), Skewness, Kurtosis, Number of Zero Crossings, Hjorth mobility and complexity, Fractal dimensions, Entropies, and the Power in Different Energy Bands, such as Delta.
3. An ensemble of trees using CatBoost, with weights to handle class imbalance considering the augmentation strategy, is trained to classify seizures and non-seizures in the 2-second epochs.

The test cycle commences with Step 2, and the resulting features are input to the trained CatBoost model to get the predictions for each epoch.

For T2, we used FT Surrogate for EEG augmentation and handling class imbalance between seizure and non-seizure. The relative class weights during training were updated in response to the augmentation effect.

3. EXPERIMENTS

Experimental Setup. SeizeIT1 [3] was randomly split into a training and validation set in an 8:2 ratio. Results in the

This research was funded by the National Science Foundation CAREER grant 1944247 and the National Institute of Health grant U19-AG056169 subaward to C.M.

following sections are from evaluation on SeizeIT2[1] [1].

Metrics. Sensitivity using the any-overlap method (OVLP) [9] and False Alarm (FA) rate using epoch-based scoring (EPOCH) [9] are used for evaluation. A weighting factor of -0.4 was used to balance sensitivity and FA rate, and the scores were averaged with weights of 0.6 for Task 1 and 0.4 for Task 2. Details are provided in the challenge website [1].

$$\text{Points}_\text{x} = \text{Sensitivity}_\text{x} - 0.4 * \text{FAs}_\text{x}/\text{hr} \quad (1)$$

$$\text{Total points} = 0.6 * \text{Points}_1 + 0.4 * \text{Points}_2 \quad (2)$$

Table 1: Model Evaluation on SeizeIT2 [1]

	Sensitivity OVLP [9]	False alarm per hour EPOCH [9]	Total Points
ChronoNet [4]	58.22	117.12	11.37
AttentionNet	53.57	30.85	41.23
`SeizFt` (T1)	**62.86**	14.93	**56.88**
ChronoNet (T2)	22.22	**9.82**	18.30

Results. Table 1 compares the results of the proposed approaches for the two tasks `SeizFt` (T1) and ChronoNet (T2) with a proposed deep neural network with multi-headed attention, we refer to as AttentionNet, and the baseline ChronoNet [4]. `SeizFt` (T1) outperforms the other frameworks. The effect of the augmentation and class balancing strategy is demonstrated by the performance improvement of ChronoNet [4] in Task 2 by significantly reducing the False Alarm rate and improving the Total Points.

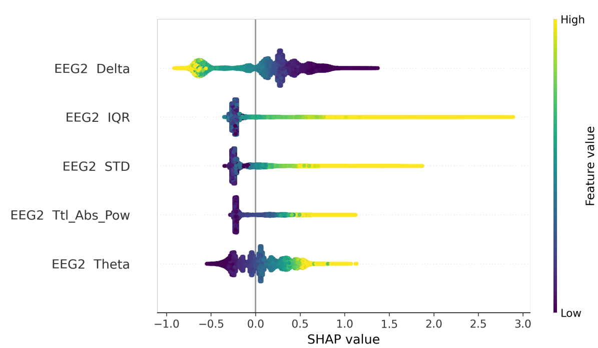

Fig. 1: Most Important Features of `SeizFt`

Interpretation. Fig. 1 highlights the five most important features in `SeizFt`. The 1st feature represents the lower delta band of 0.4-1 Hz, with lower values positively affecting seizure classification. On the other hand, the 5th feature representing the theta band of 4-8 Hz has the opposite effect. Thus, `SeizFt` shows that seizures are denoted by higher frequency EEG.

[1] `https://biomedepi.github.io/seizure_detection_challenge/`

4. CONCLUSION

The proposed approach, `SeizFt`, demonstrates the potential of robust features paired with a reliable augmentation strategy in detecting seizures, a necessity for an epilepsy diagnosis. It highlights that interpretable approaches can outperform black-box deep learning methods in the physiological monitoring of rare events.

5. REFERENCES

[1] Christos Chatzichristos, Lauren Swinnen, Jaiver Macea, Miguel Bhagubai, Wim Van Paesschen, and Maarten De Vos, "Multimodal detection of typical absence seizures in home environment with wearable electrodes," *Frontiers in Signal Processing*, vol. 2, 2022.

[2] Irfan Al-Hussaini and Cassie S Mitchell, "Serf: Interpretable sleep staging using embeddings, rules, and features," in *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 2022, pp. 3791–3795.

[3] Christos Chatzichristos and Miguel Claro Bhagubai, "SeizeIT1," 2023.

[4] Subhrajit Roy, Isabell Kiral-Kornek, and Stefan Harrer, "Chrononet: a deep recurrent neural network for abnormal eeg identification," in *Artificial Intelligence in Medicine: 17th Conference on Artificial Intelligence in Medicine, AIME 2019, Poznan, Poland, June 26–29, 2019, Proceedings 17*. Springer, 2019, pp. 47–56.

[5] Justus TC Schwabedal, John C Snyder, Ayse Cakmak, et al., "Addressing class imbalance in classification problems of noisy signals by using fourier transform surrogates," *arXiv preprint arXiv:1806.08675*, 2018.

[6] Robin Tibor Schirrmeister, Jost Tobias Springenberg, et al., "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[7] Jeroen Van Der Donckt, Jonas Van Der Donckt, et al., "Do not sleep on traditional machine learning: Simple and interpretable techniques are competitive to deep learning for sleep scoring," *Biomedical Signal Processing and Control*, vol. 81, pp. 104429, 2023.

[8] Irfan Al-Hussaini and Cassie S. Mitchell, "Performance and utility trade-off in interpretable sleep staging," in *NeurIPS 2022 Workshop on Learning from Time Series for Health*, 2022.

[9] Vinit Shah, Meysam Golmohammadi, et al., "Objective evaluation metrics for automatic classification of eeg events," *Biomedical Signal Processing: Innovation and Applications*, pp. 223–255, 2021.